Reduction of invariant constrained systems using anholonomic frames

M. Crampin and T. Mestdag
Department of Mathematics, Ghent University
Krijgslaan 281, S9, B–9000 Gent, Belgium

Abstract. We analyze two reduction methods for nonholonomic systems that are invariant under the action of a Lie group on the configuration space. Our approach for obtaining the reduced equations is entirely based on the observation that the dynamics can be represented by a second-order differential equations vector field and that in both cases the reduced dynamics can be described by expressing that vector field in terms of an appropriately chosen anholonomic frame.

Mathematics Subject Classification (2000). 34A26, 37J60, 70G45, 70H03.

Keywords. Lagrangian system, nonholonomic constraints, anholonomic frames, reduction, symmetry group

1 Introduction

In a number of recently published papers [8, 9, 10, 19] we have developed a distinctive geometric approach to the study of regular Lagrangian dynamical systems, and especially to the problem of formulating reduced equations for systems which are invariant under the action of a symmetry group. The main distinctive features of our approach are, firstly, the formulation of the Euler-Lagrange equations in a way which is well adapted to the idea that their function is to determine a vector field on the velocity phase space which is of second-order type (so that the differential equations which determine its integral curves are of second order in the configuration space coordinates), and yet is completely coordinate independent; and secondly, the consistent use of anholonomic frames and their associated quasi-velocities. In this paper we shall extend these ideas to cover Lagrangian systems subject to nonholonomic linear constraints, for which both the Lagrangian function and the constraint distribution are invariant. Such constraints arise naturally in the context of systems with rigid bodies rolling without slipping over a surface or possessing a contact point with the surface in the form of a knife edge. A classical reference for the dynamics of mechanical systems with nonholonomic constraints is the book by Neĭmark and Fufaev [21]. The recent books [1, 7, 12] contain many references to different modern geometric approaches to the theory. We will work with autonomous systems; for formulations of the nonholonomic dynamics in a time-dependent set-up see e.g. [16, 22].

The formulation of the Euler-Lagrange equations mentioned above goes as follows. We consider a Lagrangian system over a differentiable manifold Q (configuration space). The Lagrangian L is a function on the tangent bundle $\tau : TQ \to Q$ (velocity phase space); it is regular if its Hessian with respect to the fibre coordinates is nonsingular. The following proposition holds [10].

Proposition 1. *Let L be a regular Lagrangian on TQ. There is a unique second-order differential equation field Γ such that*

$$\Gamma(Z^{\mathrm{V}}(L)) - Z^{\mathrm{C}}(L) = 0$$

for all vector fields Z on Q. Moreover, Γ may be determined from the equations

$$\Gamma(X_i^{\mathrm{V}}(L)) - X_i^{\mathrm{C}}(L) = 0, \quad i = 1, 2, \dots n = \dim Q,$$

for any frame $\{X_i\}$ on Q (which may be a coordinate frame or may be anholonomic).

Here Z^{V} and Z^{C} are respectively the vertical lift and the complete or tangent lift of Z to TQ (we refer to [11] for the most common notions on tangent bundle geometry). The formulation in terms of the frame $\{X_i\}$ leads directly to Hamel's equations

$$\Gamma\left(\frac{\partial L}{\partial v^i}\right) - X_i^j \frac{\partial L}{\partial q^j} + R_{ik}^j v^k \frac{\partial L}{\partial v^j} = 0,$$

where $X_i = X_i^j \partial/\partial q^j$; the v^i are the quasi-velocities associated with the frame; and the coefficients R_{ij}^k are defined by $[X_i, X_j] = R_{ij}^k X_k$ and are collectively called the object of anholonomity of the frame. We refer the reader to Section 4 for more details.

The equations determining the dynamics of a regular system subject to nonholonomic linear constraints admit a rather similar formulation. The constraints may be specified in either of two equivalent ways: as a distribution $\mathcal{D}$ on Q (the constraint distribution), or as a submanifold $\mathcal{C}$ of TQ (the constraint submanifold). The two are related as follows: $\mathcal{C} = \{(q,u) \in TQ : u \in \mathcal{D}_q \subset T_qQ\}$. We assume that the dimension of each $\mathcal{D}_q$, and equivalently the fibre dimension of $\mathcal{C}_q$, is constant and equal to m. A vector field Γ on $\mathcal{C}$ is said to be of second-order type if it satisfies $\tau_{*(q,u)}\Gamma = u$ for all $(q,u) \in \mathcal{C}$. A Lagrangian function L is said to be regular with respect to $\mathcal{D}$ if for any local basis $\{X_\alpha\}$ of $\mathcal{D}$, $1 \le \alpha \le m$, the symmetric $m \times m$ matrix whose entries are $X_\alpha^{\mathrm{V}}(X_\beta^{\mathrm{V}}(L))$ (functions on $\mathcal{C}$) is nonsingular. In [10] we proved the following proposition.

Proposition 2. *Let L be a Lagrangian on TQ which is regular with respect to $\mathcal{D}$. Then there is a unique vector field Γ on $\mathcal{C}$ which is of second-order type, is tangent to $\mathcal{C}$, and is such that on $\mathcal{C}$*

$$\Gamma(Z^{\mathrm{V}}(L)) - Z^{\mathrm{C}}(L) = 0$$

for all $Z \in \mathcal{D}$. Moreover, Γ may be determined from the equations

$$\Gamma(X_\alpha^{\mathrm{V}}(L)) - X_\alpha^{\mathrm{C}}(L) = 0, \quad \alpha = 1, 2, \dots m,$$

on $\mathcal{C}$, where $\{X_\alpha\}$ is any local basis for $\mathcal{D}$.

This is our version of the Lagrange-d'Alembert principle (see [1, 7] for other versions); the vector field Γ is the dynamical field of the constrained system.

The formal similarity between the standard Euler-Lagrange equations and the Lagrange-d'Alembert equations in these formulations is self-evident. We shall exploit this similarity in deriving the reduced equations for an invariant constrained system: as we shall show, to obtain those equations it is enough to follow the reduction procedure for an invariant unconstrained system, while restricting attention to the constraint submanifold.

There are in fact two well-known ways of reducing the equations of an invariant unconstrained Lagrangian system. One method does not even take the Lagrangian structure of the system into account and simply involves factoring out by the action of the group, and leads to the so-called Lagrange-Poincaré equations; this is described in e.g. [5, 13, 19]. The second does take advantage of momentum conservation; its first step is to restrict to a level set of momentum, and this is followed by a reduction with respect to the invariance group of the chosen value of the momentum. This is a generalized version of Routh's procedure; it is discussed in [8, 18]. For a constrained system, however, these two methods are not equally applicable. This is because, although the constraint distribution is invariant under the symmetry group, it is not usually the case that any fundamental vector field of the action belongs to it. There is consequently no conservation of momentum, and no possiblility of Routh-type reduction. The greater part of this paper is therefore devoted to the adaptation of Lagrange-Poincaré reduction to constrained systems. This we discuss in full generality: whereas many other papers ([2, 3, 6, 20] for example) restrict their attention to the case in which at each point q of Q the constraint distribution $\mathcal{D}_q$ and the tangent space $\mathcal{V}_q$ to the orbit of the action together span T_qQ, we make no such so-called 'dimension assumption'; our only requirement is that $\mathcal{D}_q \cap \mathcal{V}_q$ has constant dimension.

Though Routh-type reduction is not possible in general, it can arise in particular cases, where there is a Lie subgroup H of the symmetry group G, necessarily normal, with Lie algebra $\mathfrak{h}$, such that for all $\xi \in \mathfrak{h}$ the corresponding fundamental vector field $\tilde{\xi}$ lies in $\mathcal{D}$. Symmetries belonging to H are said to be horizontal. We devote a separate section to the discussion of this case.

The paper is laid out as follows. In the following section we deal with the fundamental definitions and results concerning invariance of a constrained system under the free and proper action of a Lie group G on Q, leading to a version of the Atiyah sequence for such a system. In Section 3 we give a resumé in general terms of the Lagrange-Poincaré reduction procedure, and show how it may be adapted to the case of an invariant constrained system. In Section 4 we derive explicit Hamel-type formulae, in terms of a (possibly) anholonomic frame, for the Lagrange-Poincaré equations and the Lagrange-d'Alembert-Poincaré equations successively. In Section 5 we discuss Routh-type reduction for systems with horizontal symmetries.

2 Invariance of nonholonomic systems

Assume from now on that a connected Lie group G acts in a free and proper way on the left on the configuration manifold Q. Then $\pi : Q \to Q/G$ is a principal fibre bundle. The action ψ_g on Q induces an action $T\psi_g$ on TQ. We will write $\tilde{A}$ for the infinitesimal generator of the action on Q, associated to a Lie algebra element $A \in \mathfrak{g}$. Then $\tilde{A}^{\rm C}$ is an infinitesimal generator for the action on TQ. As in e.g. [3], we say that the nonholonomic system is invariant under G, or that it admits G as a symmetry group, if both the Lagrangian L and the constraint submanifold $\mathcal{C}$ of the system are invariant under the induced action of G on TQ.

Proposition 3. *The constraint submanifold $\mathcal{C} \subset TQ$ is invariant under $T\psi$ if and only if the constraint distribution $\mathcal{D}$ on Q is invariant under ψ.*

Proof. Since $\mathcal{C} = \{(q,u) : u \in \mathcal{D}_q\}$, $\mathcal{C}$ is invariant under $T\psi$ if and only if for every $q \in Q$, $u \in \mathcal{D}_q$ and $g \in G$, $\psi_{g*}u \in \mathcal{D}_{\psi_g(q)}$; that is to say, for every $q \in Q$ and $g \in G$, $\mathcal{D}_{\psi_g(q)} = \psi_{g*}\mathcal{D}_q$. □

Proposition 4. *If L is regular with respect to $\mathcal{D}$ the vector field Γ is invariant under the induced action of G on $\mathcal{C}$.*

Proof. For any $A \in \mathfrak{g}$ and $Z \in \mathcal{D}$ we have

$$\begin{aligned}
0 &= \tilde{A}^{\mathrm{C}}\big(\Gamma(Z^{\mathrm{V}}(L)) - Z^{\mathrm{C}}(L)\big) \\
&= [\tilde{A}^{\mathrm{C}}, \Gamma](Z^{\mathrm{V}}(L)) - \Gamma(\tilde{A}^{\mathrm{C}}(Z^{\mathrm{V}}(L))) - [\tilde{A}^{\mathrm{C}}, Z^{\mathrm{C}}](L) \\
&= [\tilde{A}^{\mathrm{C}}, \Gamma](Z^{\mathrm{V}}(L)) - \Gamma([\tilde{A}, Z]^{\mathrm{V}}(L)) - [\tilde{A}, Z]^{\mathrm{C}}(L).
\end{aligned}$$

Now $[\tilde{A}, Z] \in \mathcal{D}$, due to the assumed invariance of $\mathcal{D}$. By the Lagrange-d'Alembert equation the last two terms above vanish. On the other hand, the bracket $[\tilde{A}^{\mathrm{C}}, \Gamma]$ is vertical. This is certainly true for a second-order differential equation field on TQ, by a simple calculation in coordinates. Now Γ is a second-order differential equation field on $\mathcal{C}$; but we can evidently extend it to a second-order differential equation field on a neighbourhood of $\mathcal{C}$ in TQ. Since both $\tilde{A}^{\mathrm{C}}$ and Γ are tangent to $\mathcal{C}$, so also is their bracket. So on $\mathcal{C}$, $[\tilde{A}^{\mathrm{C}}, \Gamma]$ is independent of the choice of extension, and is vertical. It follows from the fact that the equation $[\tilde{A}^{\mathrm{C}}, \Gamma](Z^{\mathrm{V}}(L)) = 0$ holds for all $Z \in \mathcal{D}$, and the assumption that L is regular with respect to $\mathcal{D}$, that $[\tilde{A}^{\mathrm{C}}, \Gamma] = 0$. This may easily be seen by expressing everything in terms of the vertical lifts of a local basis for $\mathcal{D}$. But $[\tilde{A}^{\mathrm{C}}, \Gamma] = 0$ is the infinitesimal condition for Γ to be invariant. $\square$

Since Γ is invariant, it reduces to a vector field $\check{\Gamma}$ on $\mathcal{C}/G$. Our main overall aim in this paper is to show how to determine this reduced vector field. We begin however by considering some aspects of the structure of nonholonomic systems which are invariant in the sense defined above.

As we have already noted, $\pi : Q \to Q/G$ is a principal fibre bundle. Since $\mathcal{D}$ is invariant it defines a distribution $\bar{\mathcal{D}}$ on Q/G by $\bar{\mathcal{D}}_{\pi(q)} = \pi_*(\mathcal{D}_q)$; this is well-defined because $\pi \circ \psi_g = \pi$. Let us assume that $\bar{\mathcal{D}}$ has constant dimension. Then Q/G is equipped with a regular distribution $\bar{\mathcal{D}}$. Denote the corresponding submanifold (indeed vector subbundle) of $T(Q/G)$ by $\bar{\mathcal{C}}$.

Let $\mathcal{V}_q = \ker \pi_{*q}$. Note that $\ker \pi_{*q}|_{\mathcal{D}_q} = \mathcal{D}_q \cap \mathcal{V}_q$. Let us denote it by $\mathcal{S}_q$. Evidently $\mathcal{S}$ is an invariant distribution on Q, which is of constant dimension by the corresponding assumption for $\bar{\mathcal{D}}$. Since $\mathcal{S}_q \subset \mathcal{V}_q$ for each $q \in Q$, we may identify $\mathcal{S}_q$ with a vector subspace $\mathfrak{g}^q$ of $\mathfrak{g}$, where $\mathfrak{g}^q = \{A \in \mathfrak{g} \,|\, \tilde{A}_q \in \mathcal{S}_q\}$. In terms of TQ, we can express $\mathfrak{g}^q$ as follows. For $w \in T_qQ$, $w \in \mathcal{D}_q$ if and only if w^{V} is tangent to $\mathcal{C}$; thus $A \in \mathfrak{g}^q$ if and only if $\tilde{A}_q^{\mathrm{V}}$ is tangent to $\mathcal{C}$. Since (see e.g. [11, 17])

$$\psi_{g*}\left(\tilde{A}_q\right) = \left(\widetilde{\mathrm{ad}(g^{-1})A}\right)_{\psi_g(q)}$$

we have

$$\mathfrak{g}^{\psi_g(q)} = \mathrm{ad}(g^{-1})\mathfrak{g}^q.$$

Consider $\mathfrak{g}^{\mathcal{D}} = \{(q, A) \,|\, A \in \mathfrak{g}^q\}$. There is an action of G on $\mathfrak{g}^{\mathcal{D}}$ given by

$$(q, A) \mapsto (\psi_g(q), \mathrm{ad}(g^{-1})A).$$

On taking the quotient we obtain a vector bundle over Q/G, say $\bar{g}^{\mathcal{D}}$, which is a vector subbundle of $\bar{\mathfrak{g}} = (Q \times \mathfrak{g})/G \to Q/G$, the adjoint bundle associated with the principal G-bundle Q.

Proposition 5. *We have the following short exact sequence of vector bundles over Q/G:*

$$0 \to \bar{\mathfrak{g}}^{\mathcal{D}} \to \mathcal{C}/G \to \bar{\mathcal{C}} \to 0.$$

This is a version for constrained systems of the so-called Atiyah sequence (see e.g. [9, 13]),

$$0 \to \bar{\mathfrak{g}} \to TQ/G \to T(Q/G) \to 0.$$

Each term in the sequence of the proposition is a subbundle of the corresponding term in the Atiyah sequence. In the next section we will use this observation when we divide the reduced equations for an invariant nonholonomic system into two sets.

3 Lagrange-Poincaré reduction: generalities

3.1 Standard Lagrange-Poincaré reduction

Before considering reduction of invariant nonholonomic systems we discuss Lagrange-Poincaré reduction of the standard Euler-Lagrange equations. Recall the Euler-Lagrange equations as they appear in Proposition 1. Assume that L is G-invariant: then so is Γ; it reduces to a vector field $\check{\Gamma}$ on TQ/G, which we want equations for — so-called reduced equations.

There is a sense in which the reduction of the Euler-Lagrange equations is immediate. The key step is to rewrite them in G-invariant form. It is enough to take Z to be invariant. Then Z^{V} and Z^{C} are invariant under the induced action of G on TQ, and so define vector fields on TQ/G, which we denote by $\check{Z}^{\mathrm{V}}$ and $\check{Z}^{\mathrm{C}}$, though of course they are not vertical or complete lifts. The function $Z^{\mathrm{C}}(L)$ is invariant, and so defines a function on TQ/G, which is just $\check{Z}^{\mathrm{C}}(l)$ (where l is the reduced function of L on TQ/G); likewise for $Z^{\mathrm{V}}(L)$. Then the reduced equations are simply

$$\check{\Gamma}(\check{Z}^{\mathrm{V}}(l)) - \check{Z}^{\mathrm{C}}(l) = 0,$$

on TQ/G, for all invariant vector fields Z on Q; they are called the Lagrange-Poincaré equations.

However, we can be more explicit. The Euler-Lagrange equations can be divided into two sets, according to whether we take Z to be tangent to the fibres of π or transverse to them.

The Lagrange-Poincaré equation for momentum. Consider first the Euler-Lagrange equation $\Gamma(Z^{\mathrm{V}}(L)) - Z^{\mathrm{C}}(L) = 0$ for any vector field Z which is vertical with respect to $\pi : Q \to Q/G$. Such a vector field is determined by a $\mathfrak{g}$-valued function ζ on Q, where $Z_q = \widetilde{\zeta(q)}_q$. The momentum p is a $\mathfrak{g}^*$-valued function on TQ, which is G-equivariant under the usual action of G on TQ and the coadjoint action on $\mathfrak{g}^*$.

Take first $A \in \mathfrak{g}$. We have

$$\tilde{A}^{\mathrm{V}}(L) = \langle A, p \rangle$$

(as real-valued functions on TQ; the angle brackets denote the pairing of $\mathfrak{g}$ and $\mathfrak{g}^*$). The conservation of momentum is just $\Gamma\langle A, p\rangle = 0$ (the Euler-Lagrange equation with $Z = \tilde{A}$): or, since A is constant and arbitrary, $\Gamma(p) = 0$.

Now consider $\langle \zeta, p \rangle$: applying Leibniz' rule we have

$$\Gamma\langle \zeta, p\rangle = \langle \Gamma(\zeta), p\rangle + \langle \zeta, \Gamma(p)\rangle = \langle \dot{\zeta}, p\rangle$$

(using the fact that $\Gamma(f) = \dot{f}$ for any function f on Q). We claim that the (almost tautological) equation $\Gamma\langle \zeta, p\rangle = \langle \dot{\zeta}, p\rangle$ is the Euler-Lagrange equation for vertical Z. We compute Z^{V} and Z^{C} in terms of ζ, as follows. Take a basis $\{E_r\}$ of $\mathfrak{g}$, and set $\zeta = \zeta^r E_r$, where the coefficients ζ^r are functions on Q. Then $Z = \zeta^r \tilde{E}_r$, and so

$$Z^{\mathrm{V}} = \zeta^r \tilde{E}_r^{\mathrm{V}}, \quad Z^{\mathrm{C}} = \zeta^r \tilde{E}_r^{\mathrm{C}} + \dot{\zeta}^r \tilde{E}_r^{\mathrm{V}}.$$

Then

$$Z^{\mathrm{V}}(L) = \zeta^r p_r = \langle \zeta, p\rangle, \quad Z^{\mathrm{C}}(L) = \dot{\zeta}^r p_r = \langle \dot{\zeta}, p\rangle$$

as claimed.

For reduction we need to take Z to be G-invariant: it will be so if and only if ζ is G-equivariant (with now the adjoint action on $\mathfrak{g}$); this amounts to taking a section of the adjoint bundle

$\bar{\mathfrak{g}} \to Q/G$ over Q. It is then clear that $\langle \zeta, p\rangle$ will be invariant. It must then be the case (from the Euler-Lagrange equation and the invariance of Γ) that $\langle \dot{\zeta}, p\rangle$ is invariant. This can be shown directly too, in various ways. Here is a vector field version. The G-equivariance of p can be expressed as

$$\langle B, \tilde{A}^{\mathrm{C}}(p)\rangle = -\langle [A,B], p\rangle,$$

where $A, B \in \mathfrak{g}$ and $[A,B]$ is their bracket in $\mathfrak{g}$. The assumed G-equivariance of ζ is just $\tilde{A}(\zeta) = [A, \zeta]$ (again, bracket in $\mathfrak{g}$). Then

$$\begin{aligned}
\tilde{A}^{\mathrm{C}}\langle \dot{\zeta}, p\rangle &= \langle \tilde{A}^{\mathrm{C}}(\dot{\zeta}), p\rangle + \langle \dot{\zeta}, \tilde{A}^{\mathrm{C}}(p)\rangle \\
&= \left\langle \frac{d}{dt}(\tilde{A}(\zeta)), p \right\rangle - \langle [A, \dot{\zeta}], p\rangle \\
&= \left\langle \frac{d}{dt}([A, \zeta]), p \right\rangle - \langle [A, \dot{\zeta}], p\rangle = 0,
\end{aligned}$$

using the obvious fact that $X^{\mathrm{C}}(\dot{f}) = d/dt(X(f))$.

Since $\langle \zeta, p\rangle$ and $\langle \dot{\zeta}, p\rangle$ are invariant they define functions on TQ/G, which we denote by $\langle\!\langle \zeta, p\rangle\!\rangle$ and $\langle\!\langle \dot{\zeta}, p\rangle\!\rangle$. The corresponding reduced equation is

$$\check{\Gamma}\langle\!\langle \zeta, p\rangle\!\rangle = \langle\!\langle \dot{\zeta}, p\rangle\!\rangle.$$

We call it the Lagrange-Poincaré equation for momentum.

The horizontal Lagrange-Poincaré equation. To obtain an invariant Euler-Lagrange equation corresponding to the transverse directions we can use a principal connection on $\pi : Q \to Q/G$, or, equivalently, a splitting of the Atiyah sequence. Suppose we have such a connection: for any vector field Y on Q/G let Y^{H} be its horizontal lift to Q; then the transverse (let's call it horizontal) Euler-Lagrange equation is

$$\Gamma((Y^{\mathrm{H}})^{\mathrm{V}}(L)) - (Y^{\mathrm{H}})^{\mathrm{C}}(L) = 0.$$

Incidentally, this expression is $C^\infty(Q/G)$-linear in Y. Moreover, each term is G-invariant.

We will express this equation in a different way. First we recall the construction of the Vilms connection from [9] (which is a special case of a more general construction in [23]). The complete lift of a type $(1,1)$ tensor T on Q is given by [11]

$$T^{\mathrm{C}}(X^{\mathrm{C}}) = T(X)^{\mathrm{C}}, \quad T^{\mathrm{C}}(X^{\mathrm{V}}) = T(X)^{\mathrm{V}}.$$

The original connection on $\pi : Q \to Q/G$ can be represented by a type $(1,1)$ tensor field ω, so that $\omega(X) = 0$ if and only if X is horizontal, $\omega(V) = V$ for V vertical. Then the type $(1,1)$ tensor on TQ defining the Vilms connection is just ω^{C}; moreover, it is invariant under the action of G on TQ. From the defining relations of the complete lift of a type $(1,1)$ tensor field above, one easily concludes that the horizontal distribution defined by the Vilms connection is spanned by the complete and vertical lifts of the horizontal vector fields of the original connection.

Next, a remark about complete and vertical lifts. Let $\phi : M \to N$ be a smooth map, and suppose that vector fields U on M and V on N are ϕ-related. Then U^{C} and V^{C} are $T\phi$-related, and likewise U^{V} and V^{V}. One can easily prove this by considering the flows of the involved vector fields.

Consider now $(X^{\mathrm{H}})^{\mathrm{V}}$, for any vector field X on Q/G. It is horizontal with respect to the Vilms connection. By the previous remark, since X^{H} is π-related to X, $(X^{\mathrm{H}})^{\mathrm{V}}$ is $T\pi$-related to X^{V}.

Thus $(X^{\rm H})^{\rm V}$ is the horizontal lift with respect to the Vilms connection of the vector field $X^{\rm V}$ on $T(Q/G)$. With some abuse of notation we may write $(X^{\rm H})^{\rm V} = (X^{\rm V})^{\rm H}$ (warning: V and H have different meanings either side of the equality sign). Likewise for $X^{\rm C}$. So we can rewrite the horizontal Euler-Lagrange equation above as follows:

$$\Gamma((Y^{\rm V})^{\rm H}(L)) - (Y^{\rm C})^{\rm H}(L) = 0$$

for all Y on Q/G.

The Vilms connection is G-invariant, so each of $(Y^{\rm V})^{\rm H}$ and $(Y^{\rm C})^{\rm H}$ is an invariant vector field on TQ, and so each passes to the quotient to define vector fields $(Y^{\rm V})^{\check{\rm H}}$ and $(Y^{\rm C})^{\check{\rm H}}$ on TQ/G. The reduced horizontal Euler-Lagrange equation is

$$\check{\Gamma}((Y^{\rm V})^{\check{\rm H}}(l)) - (Y^{\rm C})^{\check{\rm H}}(l) = 0$$

on TQ/G, for all Y on Q/G. We call it the horizontal Lagrange-Poincaré equation.

In the next step, we can write this equation as

$$\check{\Gamma}(\langle Y^{\rm V}, d^{\check{\rm H}}l\rangle) - \langle Y^{\rm C}, d^{\check{\rm H}}l\rangle = 0.$$

Here $d^{\check{\rm H}}l$ is a 1-form along the projection $TQ/G \to T(Q/G)$ (or a 1-form on TQ/G which is semi-basic with respect to that projection), such that for any vector field W on $T(Q/G)$, $\langle W, d^{\check{\rm H}}l\rangle = W^{\check{\rm H}}(l)$. Since $Y^{\rm V}$ and $Y^{\rm C}$ are actually the lifts from Q/G to $T(Q/G)$ this looks very much like an Euler-Lagrange equation on $T(Q/G)$ (in fact, it would be one if $d^{\check{\rm H}}l$ was replaced by the exterior derivative d on $T(Q/G)$).

We have divided the reduced equations in two sets, in accordance with the decomposition of TQ/G given by the Atiyah sequence. We conclude therefore:

Proposition 6. *The Lagrange-Poincaré equations are given by*

$$\begin{aligned}&\check{\Gamma}\langle\!\langle \zeta, p\rangle\!\rangle = \langle\!\langle \dot{\zeta}, p\rangle\!\rangle\\ &\check{\Gamma}(\langle Y^{\rm V}, d^{\check{\rm H}}l\rangle) - \langle Y^{\rm C}, d^{\check{\rm H}}l\rangle = 0,\end{aligned}$$

where ζ is any G-equivariant $\mathfrak{g}$-valued function on TQ and Y is any vector field on Q/G.

3.2 Lagrange-Poincaré-type reduction of nonholonomic systems

Again, there is a sense in which the reduction of the Lagrange-d'Alembert equations is immediate. These equations say that on $\mathcal{C}$

$$\Gamma(Z^{\rm V}(L)) - Z^{\rm C}(L) = 0$$

for all $Z \in \mathcal{D}$. We assume that $\mathcal{D}$ is G-invariant. It is again enough to take Z to be invariant. The function $Z^{\rm C}(L)$ is invariant, and so defines a function on TQ/G, which is just $\check{Z}^{\rm C}(l)$ (where l is the reduced function of L). Likewise for $Z^{\rm V}(L)$; however, since $Z \in \mathcal{D}$, $Z^{\rm V}$ is tangent to $\mathcal{C}$ and so we can replace l by l_c in the first term. Then the reduced equation is simply

$$\check{\Gamma}(\check{Z}^{\rm V}(l_c)) - \check{Z}^{\rm C}(l) = 0$$

on $\mathcal{C}/G$, for all $Z \in \mathcal{D}$; of course the second term must be understood as the restriction of that function on TQ/G to the submanifold $\mathcal{C}/G$.

The reduced equations are now called the Lagrange-d'Alembert-Poincaré equations. Like the Lagrange-Poincaré equations they can be split into two sets, corresponding to the vertical and horizontal parts of $\mathcal{D}$. Much as before, these two sets are dictated by the version of the Atiyah sequence we found in Proposition 5.

The Lagrange-d'Alembert-Poincaré equation for momentum. First, we consider vertical vector fields in $\mathcal{D}$, that is, vector fields Z in $\mathcal{S}$. Every such vector field defines a $\mathfrak{g}$-valued function ζ on Q, where now $\zeta(q) \in \mathfrak{g}^q$. Since L is G-invariant we may define the momentum map $p : TQ \to \mathfrak{g}^*$, as usual, and it is G-equivariant with respect to the coadjoint action on $\mathfrak{g}^*$. However, we now have no reason to suppose that in general p, or any component of it, is conserved. On the other hand, the argument that leads to the formulae $Z^{\mathrm{V}}(L) = \langle \zeta, p \rangle$ and $Z^{\mathrm{C}}(L) = \langle \dot{\zeta}, p \rangle$ still holds good, and we conclude that the Lagrange-d'Alembert equation for $Z \in \mathcal{S}$ can be written

$$\Gamma\langle \zeta, p \rangle = \langle \dot{\zeta}, p \rangle$$

on $\mathcal{C}$. We conclude further that the following weakened version of conservation of momentum for a constrained system holds: $\langle \zeta, \Gamma(p) \rangle = 0$ for all $\mathfrak{g}$-valued functions ζ such that $\zeta(q) \in \mathfrak{g}^q$; that is to say, for all $(q, u) \in \mathcal{C}$, $\Gamma_{(q,u)}(p) \in (\mathfrak{g}^q)^\perp$. Of course, if it should happen that for some ζ, $\langle \dot{\zeta}, p \rangle = 0$ on $\mathcal{C}$ then $\langle \zeta, p \rangle$ will be conserved. In particular, this will occur if $\mathcal{S}$ contains a fundamental vector field of the action, that is, if there is some $A \in \mathfrak{g}$ such that $A \in \mathfrak{g}^q$ for all $q \in Q$: then $\langle A, p \rangle$ (the A-component of momentum) will be conserved.

To obtain a G-invariant vector field $Z \in \mathcal{S}$ we must take ζ to be G-equivariant under the adjoint action on $\mathfrak{g}$. Then we have the reduced equation

$$\check{\Gamma}\langle\!\langle \zeta, p \rangle\!\rangle = \langle\!\langle \dot{\zeta}, p \rangle\!\rangle$$

on $\mathcal{C}/G$, where $\zeta(q) \in \mathfrak{g}^q$.

The horizontal Lagrange-d'Alembert-Poincaré equation. To obtain the reduced equation corresponding to the horizontal part of $\mathcal{D}$ we need a splitting of the modified Atiyah sequence of Proposition 5. One may derive such a splitting from a principal connection on $\pi : Q \to Q/G$ with the property that the horizontal lift of $\bar{\mathcal{D}}$ is contained in $\mathcal{D}$. Such a connection can be constructed locally, by defining its horizontal subspaces as follows. Take a local section of π. For every q in the image of the section choose some complement to $\mathcal{S}_q$ in $\mathcal{D}_q$ and extend it to a complement of $\mathcal{V}_q$ in T_qQ, smoothly over the section. Finally, extend the result along the fibres by the action of G.

The reduced equation is

$$\check{\Gamma}(\langle Y^{\mathrm{V}}, d^{\check{\mathrm{H}}} l \rangle) - \langle Y^{\mathrm{C}}, d^{\check{\mathrm{H}}} l \rangle = 0$$

as before, but now with $Y \in \bar{\mathcal{D}}$.

The conclusion of this section is therefore:

Proposition 7. *The Lagrange-d'Alembert-Poincaré equations are given by*

$$\check{\Gamma}\langle\!\langle \zeta, p \rangle\!\rangle = \langle\!\langle \dot{\zeta}, p \rangle\!\rangle$$
$$\check{\Gamma}(\langle Y^{\mathrm{V}}, d^{\check{\mathrm{H}}} l \rangle) - \langle Y^{\mathrm{C}}, d^{\check{\mathrm{H}}} l \rangle = 0$$

on $\mathcal{C}/G$, where $\zeta(q) \in \mathfrak{g}^q$ and $Y \in \bar{\mathcal{D}}$.

4 Lagrange-Poincaré-type reduction: formulae

The versions of the reduced equations given in the previous section are elegant and instructive. In the literature one may find other geometric approaches to obtain the reduced equations (see e.g. [5, 13] for the Lagrange-Poincaré equations and [1, 3, 6, 20] for the Lagrange-d'Alembert-Poincaré equations). In the interest of comparison, we shall now formulate local versions of our reduced equations. Our method is entirely based on the use of suitably adapted anholonomic frames on Q. Unsurprisingly, the version of the Lagrange-d'Alembert-Poincaré equations we finally obtain will combine elements of both the Lagrange-d'Alembert and the Lagrange-Poincaré equations, so we first deal with each of those cases separately.

But even before doing so, it is worth recalling the basic formulae relating to anholonomic frames. Let $\{X_i\}$ be an anholonomic frame on Q and v^i the quasi-velocities corresponding to that frame. (The quasi-velocities v^i are not to be confused with the canonical fibre coordinates associated with the x^i; the coordinates (x^i, v^i) are to that extent unnatural.) A second-order field Γ can then be written in the form

$$\Gamma = v^i X_i^{\mathrm{C}} + f^i X_i^{\mathrm{V}}.$$

We write $[X_i, X_j] = R_{ij}^k X_k$, where the functions R_{ij}^k are collectively called the object of anholonomity. Let v^i be the quasi-velocities corresponding to the frame: then

$$X_i^{\mathrm{C}}(v^j) = -R_{ik}^j v^k, \qquad X_i^{\mathrm{V}}(v^j) = \delta_i^j.$$

In terms of coordinates x^i on Q we may write

$$X_i^{\mathrm{C}} = X_i^j \frac{\partial}{\partial x^j} - R_{ik}^j v^k \frac{\partial}{\partial v^j}, \qquad X_i^{\mathrm{V}} = \frac{\partial}{\partial v^i},$$

where $X_i = X_i^j \partial/\partial x^j$. The first term in the expression for X_i^{C} is formally the same as X_i itself, but is of course a local vector field on TQ rather than a vector field on Q; we shall continue to denote it by X_i, though this is strictly speaking an abuse of notation. With this understood, the Euler-Lagrange equations may be written in Hamel form,

$$\Gamma\left(\frac{\partial L}{\partial v^i}\right) - X_i(L) + R_{ik}^j v^k \frac{\partial L}{\partial v^j} = 0$$

(as in e.g. [3]), which is the form we had announced in our Introduction.

4.1 The Lagrange-Poincaré equations

We take the frame $\{X_i\}$ on Q to be invariant under the action of G, and to be of the form $\{X_r, X_I\}$ where the X_r are vertical and such that their values at any point q form a basis for the vertical vectors at q. Then the X_I are invariant and transverse to the fibres of $Q \to Q/G$, and may be considered as the horizontal lifts of their projections Y_I to Q/G, with respect to some principal connection ω; note that $\{Y_I\}$ is a frame for Q/G, in general anholonomic.

Let $\{E_r\}$ be a basis of $\mathfrak{g}$ and let $\tilde{E}_r$ be the fundamental vector fields of the action corresponding to this basis. We have $[\tilde{E}_r, \tilde{E}_s] = -C_{rs}^t \tilde{E}_t$ where the coefficients C_{rs}^t are the structure constants of $\mathfrak{g}$ with respect to the given basis. A vector field on Q is invariant if and only if all $[X, \tilde{E}_r] = 0$.

The vector fields X_r could be obtained by taking a local section of $Q \to Q/G$, choosing a basis of vertical vectors at each point of the section varying smoothly over it, and using the G-action

to define the vector fields along the fibres. If one chooses the initial values of the X_r to be $\tilde{E}_r$ we obtain the vector fields $\hat{E}_r$ that we have used in previous publications [9, 19]. However, in view of what we need for the next section on constrained systems, it will be convenient to work in greater generality already in this part of the section.

We may write $X_r = X_r^s \tilde{E}_s$ where the coefficient matrix is nonsingular. We have

$$[\tilde{E}_r, X_s] = [\tilde{E}_r, X_s^t \tilde{E}_t] = (\tilde{E}_r(X_s^t) - C_{ru}^t X_s^u)\tilde{E}_t,$$

and so the necessary and sufficient condition for the X_r to be invariant is that the coefficients X_s^t satisfy

$$\tilde{E}_r(X_s^t) = C_{ru}^t X_s^u.$$

It follows immediately that $X_r(X_s^t) = C_{uv}^t X_r^u X_s^v$, whence

$$[X_r, X_s] = [X_r, X_s^t \tilde{E}_t] = C_{uv}^t X_r^u X_s^v \tilde{E}_t = (\bar{X}_w^t C_{uv}^w X_r^u X_s^v) X_t$$

where the overbar indicates the matrix inverse. Since X_r, X_s and X_t are all invariant, so must the coefficient be. We set $\bar{C}_{rs}^t = \bar{X}_w^t C_{uv}^w X_r^u X_s^v$. Then each $\bar{C}_{rs}^t$ may be treated as a function on Q/G, and the collection of such functions may be regarded as the structure constants of $\mathfrak{g}$, though expressed in terms of the X_r. (If $X_r = \hat{E}_r$ then $\bar{C}_{rs}^t = C_{rs}^t$.)

The equation $\tilde{E}_r(X_s^t) = C_{ru}^t X_s^u$ expresses the fact that the $\mathfrak{g}$-valued function $X_s^t E_t = \xi_s$ on Q is G-equivariant with respect to the adjoint action, and therefore corresponds to a section of the adjoint bundle $\bar{\mathfrak{g}} \to Q/G$; thus the ξ_r together form a local basis of sections of the adjoint bundle. Now

$$X_I(\xi_r) = X_I(X_r^s)E_s = \bar{X}_t^s X_I(X_r^t)\xi_s;$$

for convenience we shall write Υ_{Ir}^s for $\bar{X}_t^s X_I(X_r^t)$. We know that the functions Υ_{Ir}^s are G-invariant, and may therefore be regarded as functions on Q/G. Now X_I is the horizontal lift of the vector field Y_I on Q/G to Q, so we have

$$\Upsilon_{Ir}^s \xi_s = Y_I^{\mathrm{H}}(\xi_r).$$

This means that the Υ_{Ir}^s are the connection coefficients, with respect to the local basis $\{Y_I\}$ of vector fields on Q/G and the local basis $\{\xi_r\}$ of sections of $\bar{\mathfrak{g}} \to Q/G$, of the connection induced by ω on the adjoint bundle (see e.g. [9] for more details).

Since the elements of the frame $\{X_i\}$ are invariant, so are their brackets, and so is each R_{ij}^k; it may therefore be regarded as a function on Q/G. The following facts about the R_{ij}^k are important.

- The R_{IJ}^K constitute the object of anholonomity of the frame $\{Y_I\}$.
- The vertical component of $[X_I, X_J]$, which is $R_{IJ}^r X_r$, is closely related to the curvature of the connection ω: in fact the curvature, as a $\mathfrak{g}$-valued function, is $-R_{IJ}^s X_s^r E_r$. We write $-K_{IJ}^r$ for R_{IJ}^r as a reminder of this fact.
- Since $[X_i, X_r]$ is always vertical, $R_{ir}^I = 0$.
- $R_{Ir}^s = \Upsilon_{Ir}^s$.
- $R_{rs}^t = \bar{C}_{rs}^t$.

We let (v^r, v^I) be the the quasi-velocities corresponding to the frame $\{X_r, X_I\}$; thus for $v \in T_q Q$, $v = v^r X_r(q) + v^I X_I(q)$. But then $\pi_* v \in T_{\pi(q)}(Q/G)$ is given by

$$\pi_* v = v^I \pi_* X_I(q) = v^I Y_I(\pi(q)).$$

From the invariance of the frame $\{X_i\}$ we conclude that the v^i are invariant, and therefore constitute fibre coordinates on $TQ/G \to Q/G$. Since $\pi_* : T_q Q \to T_{\pi(q)}(Q/G)$ is surjective, we can identify v^I with $\pi^* w^I$, where the w^I are the quasi-velocities of the frame $\{Y_I\}$.

We now consider the Euler-Lagrange equation for momentum, $\Gamma(X_r^{\mathrm{V}}(L)) - X_r^{\mathrm{C}}(L) = 0$. The components of momentum are of course given by $p_r = \tilde{E}_r^{\mathrm{V}}(L)$. But we are working in terms of the invariant basis $\{X_r\}$. Let us set $P_r = X_r^{\mathrm{V}}(L) = X_r^s p_s = \langle \xi_r, p \rangle$; then P_r is invariant. The Euler-Lagrange equation becomes

$$\Gamma(P_r) = \dot{X}_r^s p_s = \bar{X}_t^s \dot{X}_r^t P_s,$$

which is the component form of the momentum equation given in Section 3, and is usually referred to as the momentum equation in a moving basis [1, 2]. Now in terms of quasi-velocities

$$\frac{d}{dt} = v^i X_i = v^I X_I + v^r X_r,$$

whence

$$\bar{X}_t^s \dot{X}_r^t = \bar{X}_t^s (v^I X_I(X_r^t) + v^u X_u(X_r^t)) = \Upsilon_{Ir}^s v^I - \bar{C}_{rt}^s v^t,$$

and we have

$$\Gamma(P_r) = (\Upsilon_{Ir}^s v^I - \bar{C}_{rt}^s v^t) P_s.$$

Taking account of the known facts about the R_{ij}^k, together with the invariance of the frame, we have

$$\begin{aligned}
\check{X}_r^{\mathrm{C}} &= -R_{ri}^s v^i \frac{\partial}{\partial v^s} = \left(\Upsilon_{Ir}^s v^I - \bar{C}_{rt}^s v^t\right) \frac{\partial}{\partial v^s} \\
\check{X}_r^{\mathrm{V}} &= \frac{\partial}{\partial v^r} \\
\check{X}_I^{\mathrm{C}} &= Y_I - R_{IJ}^K v^J \frac{\partial}{\partial v^K} - R_{Ij}^r v^j \frac{\partial}{\partial v^r} \\
&= Y_I - R_{IJ}^K v^J \frac{\partial}{\partial v^K} + \left(K_{IJ}^r v^J - \Upsilon_{Is}^r v^s\right) \frac{\partial}{\partial v^r} \\
\check{X}_I^{\mathrm{V}} &= \frac{\partial}{\partial v^I}.
\end{aligned}$$

By substituting these expression in the reduced equations $\check{\Gamma}(\check{X}_i^{\mathrm{V}}(l)) - \check{X}_i^{\mathrm{C}}(l) = 0$ (with $i = r$ and $i = I$ successively) we get:

Proposition 8. *The Lagrange-Poincaré equations are given by*

$$\begin{aligned}
\check{\Gamma}\left(\frac{\partial l}{\partial v^r}\right) &= \left(\Upsilon_{Ir}^s v^I - \bar{C}_{rt}^s v^t\right) \frac{\partial l}{\partial v^s} \\
\check{\Gamma}\left(\frac{\partial l}{\partial v^I}\right) - Y_I(l) + R_{IJ}^K v^J \frac{\partial l}{\partial v^K} &= \left(K_{IJ}^r v^J - \Upsilon_{Is}^r v^s\right) \frac{\partial l}{\partial v^r}.
\end{aligned}$$

The first of these is of course just the reduced form of the momentum equation given earlier.

If we take the Y_I to be coordinate fields the horizontal equation takes on a somewhat more familiar appearance:

$$\check{\Gamma}\left(\frac{\partial l}{\partial v^I}\right) - \frac{\partial l}{\partial x^I} = \left(K^r_{IJ}v^J - \Upsilon^r_{Is}v^s\right)\frac{\partial l}{\partial v^r};$$

now the v^I are effectively the standard fibre coordinates on $T(Q/G)$ (quasi no longer).

Finally, we reconcile this version of the horizontal Lagrange-Poincaré equation (for any frame $\{Y_I\}$ on Q/G) with the more abstract one given earlier,

$$\check{\Gamma}(\langle Y^{\mathrm{V}}, d^{\check{\mathrm{H}}}l\rangle) - \langle Y^{\mathrm{C}}, d^{\check{\mathrm{H}}}l\rangle = 0,$$

by computing $d^{\check{\mathrm{H}}}l$ in terms of the frame $\{Y_I\}$. Let $\{\vartheta^I\}$ be the basis of 1-forms on Q/G dual to the Y_I, and v^I the quasi-velocities. Then $\{\vartheta^I, dv^I\}$ is a basis of 1-forms on $T(Q/G)$ (we haven't distinguished notationally between 1-forms on Q/G and their pullbacks to $T(Q/G)$). Note that

$$\langle Y_I^{\mathrm{C}}, \vartheta^J\rangle = \delta^J_I, \quad \langle Y_I^{\mathrm{C}}, dv^J\rangle = -R^j_{IK}v^K, \quad \langle Y_I^{\mathrm{V}}, \vartheta^J\rangle = 0, \quad \langle Y_I^{\mathrm{V}}, dv^J\rangle = \delta^J_I.$$

Now $X_I = Y_I^{\mathrm{H}}$, from which it follows that

$$\begin{aligned}(Y_I^{\mathrm{C}})^{\check{\mathrm{H}}} &= \check{X}_I^{\mathrm{C}} = Y_I - R^K_{IJ}v^J\frac{\partial}{\partial v^K} + \left(K^r_{IJ}v^J - \Upsilon^r_{Is}v^s\right)\frac{\partial}{\partial v^r}\\ (Y_I^{\mathrm{V}})^{\check{\mathrm{H}}} &= \check{X}_I^{\mathrm{V}} = \frac{\partial}{\partial v^I}.\end{aligned}$$

Recall that $d^{\check{\mathrm{H}}}l$ is a 1-form along the projection $(TQ)/G \to T(Q/G)$, which means that it may be expressed as a linear combination of the forms $\{\vartheta^I, dv^I\}$ with coefficients which are functions on $(TQ)/G$. Using the expressions above for $(Y_I^{\mathrm{C}})^{\check{\mathrm{H}}}$ and $(Y_I^{\mathrm{V}})^{\check{\mathrm{H}}}$ we obtain

$$d^{\check{\mathrm{H}}}l = \left(Y_I(l) + \left(K^r_{IJ}v^J - \Upsilon^r_{Is}v^s\right)\frac{\partial l}{\partial v^r}\right)\vartheta^I + \frac{\partial l}{\partial v^I}dv^I,$$

which leads to the expressions for the reduced equations given above.

4.2 The Lagrange-d'Alembert-Poincaré equations

We now choose the anholonomic frame $\{X_i\}$ in the form $\{X_\alpha, X_a\}$ where $\{X_\alpha\}$ is a local basis of the distribution $\mathcal{D}$. We write the quasi-velocities as (v^α, v^a). The constraint submanifold $\mathcal{C}$ is then simply given by $v^a = 0$. The Lagrange-d'Alembert equations are $\Gamma(X_\alpha^{\mathrm{V}}(L)) - X_\alpha^{\mathrm{C}}(L) = 0$ on $\mathcal{C}$. Recall that Γ represents here a vector field on $\mathcal{C}$ of second-order type, which means that it is of the form

$$\Gamma = v^\alpha X_\alpha^{\mathrm{C}} + f^\alpha X_\alpha^{\mathrm{V}}.$$

The Lagrange-d'Alembert equations become

$$\Gamma\left(\frac{\partial L}{\partial v^\alpha}\right) - X_\alpha(L) + R^i_{\alpha\beta}v^\beta\frac{\partial L}{\partial v^i} = 0$$

in Hamel form. It is sometimes considered preferable to separate out those terms which involve differentiation along $\mathcal{C}$ from those which involve differentiation transverse to it; in the former we

can replace L by L_c, the constrained Lagrangian, in other words the restriction of L to $\mathcal{C}$. We obtain

$$\Gamma\left(\frac{\partial L_c}{\partial v^\alpha}\right) - X_\alpha(L_c) + R^\beta_{\alpha\gamma} v^\gamma \frac{\partial L_c}{\partial v^\beta} = -\left. R^a_{\alpha\beta} v^\beta \frac{\partial L}{\partial v^a}\right|_{\mathcal{C}}.$$

To obtain reduced equations for an invariant constrained system we need an adapted frame $\{X_i\} = \{X_\alpha, X_a\}$ which is invariant, as before. The basis $\{X_\alpha\}$ of $\mathcal{D}$, in turn, is of the form $\{X_\rho, X_\kappa\}$ where $\{X_\rho\}$ is a basis for $\mathcal{S}$. The set $\{X_a\}$ takes the form $\{X_c, X_k\}$ where the X_c are vertical. The collection $\{X_\rho, X_c\}$ is a basis $\{X_r\}$ of the vertical vector fields; in general we can no longer take $\hat{E}_r$ for X_r. The collection $\{X_\kappa, X_k\} = \{X_I\}$ is transverse to the fibres of $Q \to Q/G$ and is invariant, so can be taken to be the horizontal lifts of their projections Y_I to Q/G with respect to some suitable principal connection ω. The vector fields Y_κ form a basis for $\bar{\mathcal{D}}$.

The corresponding quasi-velocities are (v^α, v^a) or $(v^\rho, v^\kappa, v^c, v^k)$; the constraint submanifold $\mathcal{C}$ is given by $v^a = 0$, and (v^κ, v^k) are quasi-velocities on Q/G corresponding to the frame $\{Y_I\} = \{Y_\kappa, Y_k\}$, with $v^k = 0$ defining the constraint submanifold $\bar{\mathcal{C}}$.

The Lagrange-d'Alembert equations reduce, taking $\alpha = \rho$ and $\alpha = \kappa$ in turn, to the following pair of equations on $\mathcal{C}/G$:

$$\check{\Gamma}\left(\frac{\partial l}{\partial v^\rho}\right) = \left(\Upsilon^r_{\kappa\rho} v^\kappa - \bar{C}^r_{\rho\sigma} v^\sigma\right) \frac{\partial l}{\partial v^r}$$
$$\check{\Gamma}\left(\frac{\partial l}{\partial v^\kappa}\right) - Y_\kappa(l) + R^I_{\kappa\lambda} v^\lambda \frac{\partial l}{\partial v^I} = \left(K^r_{\kappa\lambda} v^\lambda - \Upsilon^r_{\kappa\rho} v^\rho\right) \frac{\partial l}{\partial v^r}.$$

Now L and $\mathcal{C}$ are both invariant under G, and so the constrained Lagrangian L_c is invariant under G, and defines a function l_c on $\mathcal{C}/G$, which coincides with the restriction of l (a function on $(TQ)/G$). The function l_c is called the constrained reduced Lagrangian (but might just as well be called the reduced constrained Lagrangian). We can use this to rewrite the reduced equations.

Proposition 9. *The Lagrange-d'Alembert-Poincaré equations are given by*

$$\check{\Gamma}\left(\frac{\partial l_c}{\partial v^\rho}\right) = \left(\Upsilon^r_{\kappa\rho} v^\kappa - \bar{C}^r_{\rho\sigma} v^\sigma\right) \left.\frac{\partial l}{\partial v^r}\right|_{\mathcal{C}/G}$$
$$\check{\Gamma}\left(\frac{\partial l_c}{\partial v^\kappa}\right) - Y_\kappa(l_c) + R^\lambda_{\kappa\mu} v^\mu \frac{\partial l_c}{\partial v^\lambda}$$
$$= -R^k_{\kappa\lambda} v^\lambda \left.\frac{\partial l}{\partial v^k}\right|_{\mathcal{C}/G} + \left(K^r_{\kappa\lambda} v^\lambda - \Upsilon^r_{\kappa\rho} v^\rho\right) \left.\frac{\partial l}{\partial v^r}\right|_{\mathcal{C}/G}.$$

The first of these is the reduced momentum equation. The restriction of the momentum p to $\mathfrak{g}^{\mathcal{D}}$, that is, the map $p^{\mathcal{D}} : TQ \to (\mathfrak{g}^{\mathcal{D}})^*$ given by

$$\langle \xi, p^{\mathcal{D}}(q, u)\rangle = \tilde{\xi}^{\mathrm{V}}_q(L)(q, u) \quad \text{for } \xi \in \mathfrak{g}^q,$$

is sometimes called the nonholonomic momentum map. Its components are the G-invariant functions $P_\rho = \langle \xi_\rho, p\rangle = X^{\mathrm{V}}_\rho(L)$, where $\xi_\rho = X^r_\rho E_r$ defines a section of $\bar{\mathfrak{g}}^{\mathcal{D}} \to Q/G$. The functions P_ρ satisfy $\Gamma(P_\rho) = \bar{X}^r_s \dot{X}^s_\rho P_r$ on $\mathcal{C}$, and this reduces to the Lagrange-d'Alembert-Poincaré equation for momentum given above.

We next consider some special cases of Lagrange-Poincaré-type reduction of nonholonomic systems.

In case $T_qQ = \mathcal{D}_q + \mathcal{V}_q$ (i.e. when the so-called 'dimension assumption' is satisfied), the space $\bar{\mathcal{C}}$ is the whole of $T(Q/G)$. Furthermore, we can replace the Y_I with coordinate fields $\partial/\partial x^I$ on Q/G, and the horizontal reduced equation becomes

$$\check{\Gamma}\left(\frac{\partial l_c}{\partial v^I}\right) - \frac{\partial l_c}{\partial x^I} = \left(K^r_{IJ}v^J - \Upsilon^r_{I\rho}v^\rho\right)\left.\frac{\partial l}{\partial v^r}\right|_{\mathcal{C}/G}.$$

In case $\mathcal{S}_q = \{0\}$ the constraints are said to be purely kinematic in [7]. In this case there is no momentum equation.

Chaplygin systems (see e.g. [4, 15]) are systems which have both of the above properties. There is no momentum equation, and $\mathcal{D}$ is now the horizontal distribution $\mathcal{H}$ of a principal connection. We can therefore identify $\mathcal{C}/G$ with $T(Q/G)$. The reduced vector field is now of the form

$$\check{\Gamma} = v^I\frac{\partial}{\partial x^I} + \Gamma^I\frac{\partial}{\partial v^I},$$

i.e. it is a (true) second-order differential equation field on Q/G, and its coefficients Γ^I can be determined from the equations

$$\check{\Gamma}\left(\frac{\partial l_c}{\partial v^I}\right) - \frac{\partial l_c}{\partial x^I} = K^r_{IJ}v^J\left.\frac{\partial l}{\partial v^r}\right|_{T(Q/G)}.$$

These equations are of the form of Euler-Lagrange equations subjected to an external force of gyroscopic type. See e.g. [10] for more details on this case, in the framework of anholonomic frames.

In case $\mathcal{D} \subset \mathcal{V}$ there is no horizontal equation, and the momentum equation is just

$$\check{\Gamma}\left(\frac{\partial l_c}{\partial v^\rho}\right) = -\bar{C}^r_{\rho\sigma}v^\sigma\left.\frac{\partial l}{\partial v^r}\right|_{\mathcal{C}/G}.$$

One important special case occurs when the configuration space Q is a Lie group (that is, $Q = G$), and the constraints are linear; the reduced equations are then called the Euler-Poincaré-Suslov equations in e.g. [3, 14].

5 Routh-type reduction for systems with horizontal symmetries

We now consider the class of systems with a so-called group of horizontal symmetries [2]. For that case, one assumes that there exists a subgroup $H \subset G$, the so-called group of horizontal symmetries, such that $\tilde{A} \in \mathcal{D}$ for all $A \in \mathfrak{h}$ and $\mathcal{S}_q = \mathcal{D}_q \cap \mathcal{V}_q = \mathcal{V}^H_q = \{\tilde{A}(q) \,|\, A \in \mathfrak{h}, q \in Q\}$. Because of the property $\mathfrak{g}^{\psi_q(q)} = \mathrm{ad}(g^{-1})\mathfrak{g}^q$ that we encountered in Section 2, we get that $\mathfrak{h} = \mathrm{ad}(g^{-1})\mathfrak{h}$, meaning that $\mathfrak{h}$ is necessarily an ideal (or that H is a normal subgroup).

For systems with the above properties, one can, of course, still use the reduction procedure as described in the previous sections. There are, however, also other approaches to reduction. For example, in [7] it is shown that a version of Marsden-Weinstein reduction can be applied to this case. The goal of this section is to show that one can easily stay on the 'Lagrangian side'

and recast everything in terms of Routh reduction. We will follow the geometric approach to non-Abelian Routh reduction we have developed in [8]. (For a different approach see [18].) As before, the main observation is that all one needs to do is to choose an appropriate frame. Let's assume for simplicity that in this section $\mathcal{V}_q + \mathcal{D}_q = T_qQ$.

Let $\{X_\kappa\}$ be the invariant vector fields we had before. If $\{E_r\} = \{E_\rho, E_c\}$ is a basis of $\mathfrak{g}$ whose first members $\{E_\rho\}$ span $\mathfrak{h}$, we can use $\{X_\alpha\} = \{X_\kappa, \tilde{E}_\rho\}$ as a (now not-invariant) anholonomic frame for $\mathcal{D}$, and $\{X_\kappa, \tilde{E}_\rho, \tilde{E}_c\}$ as a complete basis of vector fields on Q (with corresponding quasi-velocities $(v^\kappa, \tilde{v}^\rho, \tilde{v}^c)$. Given that $\tilde{E}^{\rm C}_\rho(L) = 0$, the Lagrange-d'Alembert equation in the direction of $\tilde{E}_\rho$ now becomes the conservation law

$$\tilde{E}^{\rm V}_\rho(L) = \mu_\rho,$$

where $\mu = \mu_\rho E^\rho \in \mathfrak{h}^*$, where the E^ρ are part of the basis that is dual to $\{E_\rho, E_c\}$. This represents a relation on $\mathcal{C}$, not on the whole of TQ.

The remaining Lagrange-d'Alembert equations are of the form

$$\Gamma(X^{\rm V}_\kappa(L)) - X^{\rm C}_\kappa(L) = 0.$$

We will restrict these equations to a fixed level set of momentum, from now on denoted by N_μ, and rewrite them in a form that contains only vector fields that are tangent to N_μ. We will do so in two steps. It is easy to see that the vector fields

$$\bar{X}^{\rm C}_\kappa = X^{\rm C}_\kappa + \tilde{R}^r_{\kappa\lambda} v^\lambda \tilde{E}^{\rm V}_r \qquad \text{and} \qquad X^{\rm V}_\kappa$$

are tangent to $\mathcal{C}$. Here $\tilde{R}^r_{\kappa\lambda}$ stands for the component of the bracket $[X_\kappa, X_\lambda]$ along $\tilde{E}_r$. There is no contribution in $\tilde{v}^\rho$ since $[X_\kappa, \tilde{E}_\rho] = 0$. The equations then become

$$\Gamma(X^{\rm V}_\kappa(L)) - \bar{X}^{\rm C}_\kappa(L) = -\tilde{R}^r_{\kappa\lambda} v^\lambda \tilde{E}^{\rm V}_r(L)$$

on N_μ.

If we further assume the matrix $(g_{\rho\sigma}) = (\tilde{E}^{\rm V}_\rho(\tilde{E}^{\rm V}_\sigma(L)))$ to be non-singular, the momentum equations can be solved in the form $\tilde{v}^\rho = \iota^\rho$, where ι^ρ are functions of the other variables. Moreover, under that assumption we can always find vector fields $W^{\rm C}_\kappa$ and $W^{\rm V}_\kappa$, with

$$\begin{aligned} W^{\rm C}_\kappa &= \bar{X}^{\rm C}_\kappa + A^\rho_\kappa \tilde{E}^{\rm V}_\rho \\ W^{\rm V}_\kappa &= X^{\rm V}_\kappa + B^\rho_\kappa \tilde{E}^{\rm V}_\rho, \end{aligned}$$

which, as well as being tangent to $\mathcal{C}$, are also tangent to the level set N_μ (the notation may again be a bit misleading since they will not be complete or vertical lifts). Let us denote $p_\rho = \tilde{E}^{\rm V}_\rho(L)$ and let us introduce the Routhian of L as the function $\mathcal{R} = L - \tilde{v}^\rho p_\rho$. Then, on the level set (which is a part of $\mathcal{C}$)

$$\begin{aligned} W^{\rm C}_\kappa(\mathcal{R}) &= W^{\rm C}_\kappa(L) - W^{\rm C}_\kappa(\tilde{v}^\rho) p_\rho = \bar{X}^{\rm C}_\kappa(L) + A^\rho_\kappa p_\rho - \bar{X}^{\rm C}_\kappa(\tilde{v}^\rho) p_\rho - A^\rho_\kappa p_\rho \\ &= \bar{X}^{\rm C}_\kappa(L) - \bar{X}^{\rm C}_\kappa(\tilde{v}^\rho) p_\rho = \bar{X}^{\rm C}_\kappa(L) + \tilde{R}^\rho_{\kappa\beta} v^\beta \mu_\rho - \tilde{R}^r_{\kappa\lambda} v^\lambda \delta^\rho_r \mu_\rho \\ &= \bar{X}^{\rm C}_\kappa(L) + \tilde{R}^\rho_{\kappa\lambda} v^\lambda \mu_\rho - \tilde{R}^\rho_{\kappa\lambda} v^\lambda \mu_\rho = \bar{X}^{\rm C}_\kappa(L) \\ W^{\rm V}_\kappa(\mathcal{R}) &= W^{\rm V}_\kappa(L) - W^{\rm V}_\kappa(\tilde{v}^\rho) p_\rho = X^{\rm V}_\kappa(L) + B^\rho_\kappa p_\rho - B^\rho_\kappa p_\rho \\ &= X^{\rm V}_\kappa(L). \end{aligned}$$

Here, $\tilde{R}^{\rho}_{\kappa\lambda}$ stands for the component of $[X_\kappa, X_\lambda]$ along $\tilde{E}_\rho$. We have also used that $[X_\kappa, \tilde{E}_\rho] = 0$.

The vector fields $\tilde{E}^{\mathrm{C}}_\rho$ are tangent to $\mathcal{C}$. We can fix functions C^σ_ρ such that the vector fields

$$\bar{E}^{\mathrm{C}}_\rho = \tilde{E}^{\mathrm{C}}_\rho + C^\sigma_\rho \tilde{E}^{\mathrm{V}}_\sigma$$

are tangent to N_μ.

Since Γ is tangent to $p_\rho = \mu_\rho$ its restriction to this level set is of the form

$$\Gamma = v^\kappa W^{\mathrm{C}}_\kappa + \iota^\rho \bar{E}^{\mathrm{C}}_\rho + (\Gamma^\kappa \circ \iota) W^{\mathrm{V}}_\kappa .$$

The coefficient $\Gamma^\kappa \circ \iota$ can be determined from the remaining Lagrange-d'Alembert equations, which take the form

$$\Gamma(W^{\mathrm{V}}_I(\mathcal{R}^\mu)) - W^{\mathrm{C}}_I(\mathcal{R}^\mu) = -\tilde{R}^{c}_{\kappa\lambda} v^\lambda \tilde{E}^{\mathrm{V}}_c(L) - \tilde{R}^{\rho}_{\kappa\lambda} v^\lambda \mu_\rho$$

on N_μ.

We can now try to understand how to reduce this restriction of Γ. It is easy to see that the action of G on $\mathcal{C}$ can be restricted to an action of the isotropy group H_μ on the level set N_μ in $\mathcal{C}$. Indeed, we have

$$0 = A^\sigma \tilde{E}^{\mathrm{C}}_\sigma(\tilde{E}^{\mathrm{V}}_\rho(L)) = -A^\sigma C^\tau_{\sigma\rho} \tilde{E}^{\mathrm{V}}_\tau(L) = -A^\sigma C^\tau_{\sigma\rho} \mu_\tau$$

if and only if $A = A^\sigma E_\sigma \in \mathfrak{h}_\mu$. We can therefore reduce the above vector field to a vector field $\check{\Gamma}_1$ on N_μ/H_μ. This reduction method is the direct analogue of the situation for standard Routh reduction (in the absence of constraints).

But there is more. Since we know that H is normal in G, the level set $\tilde{E}^{\mathrm{V}}_\rho(L) = \mu_\rho$ has also the following behaviour

$$0 = A^r \tilde{E}^{\mathrm{C}}_r(\tilde{E}^{\mathrm{V}}_\rho(L)) = -A^r C^s_{r\rho} \tilde{E}^{\mathrm{V}}_s(L) = -A^r C^\sigma_{r\rho} \tilde{E}^{\mathrm{V}}_\sigma(L) = -A^r C^\sigma_{r\rho} \mu_\sigma$$

if and only if $A = A^r E_r \in \mathfrak{g}_\mu$. Therefore, the G-action on $\mathcal{C}$ restricts in fact to a G_μ-action on the level set N_μ. We are now in the situation of a G-invariant vector field Γ on a manifold $\mathcal{C}$, which we can restrict to a G_μ-invariant vector field on N_μ and which we can reduce to a vector field $\check{\Gamma}_2$ on N_μ/G_μ.

The link with the vector field $\check{\Gamma}_1$ of the previous paragraph is the following. Instead of doing a direct reduction by G_μ, one can perform a reduction in two stages. Indeed, it is easy to define an action of G_μ/H_μ on N_μ/H_μ (see also [7]). The vector field $\check{\Gamma}_1$ will be invariant under that action and we can therefore perform a second reduction by means of its symmetry group G_μ/H_μ.

We will not write down explicit expressions for these reduced vector fields and their corresponding differential equations. Instead, we will make the situation clear by means of a simple example.

Example. Consider the system with $L = \frac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)$ on $\mathbb{R}^3$ with constraint $\dot{z} = x\dot{x}$ (a variation on the theme of a nonholonomic particle). This example is taken from [7], but we will rephrase it in our current framework. We have $\mathcal{D} = \mathrm{span}\{\partial/\partial x + x\partial/\partial z, \partial/\partial y\}$ and, since the system is invariant under the $\mathbb{R}^2$-action given by $(r,s) \times (x,y,z) \mapsto (x, y+r, z+s)$, $\mathcal{V} = \mathrm{span}\{\partial/\partial y, \partial/\partial z\}$. Therefore, $\mathcal{S} = \mathcal{D} \cap \mathcal{V} = \mathrm{span}\{\partial/\partial y\}$. This coincides with the case where $H = \mathbb{R} \times \{0\}$, with action $(r,0) \times (x,y,z) \mapsto (x, y+r, z)$. Remark that $\mathcal{D} + \mathcal{V} = TQ$.

Quasi-velocities with respect to the given frame are $v_x = \dot{x}$, $v_y = \dot{y}$ and $v_z = \dot{z} - x\dot{x}$. The vector field $X_\kappa = X = \partial/\partial x + x\partial/\partial z$ is invariant under the G-action.

The preserved momentum is here

$$\left(\frac{\partial}{\partial y}\right)^{\rm V}(L) = \dot{y} = \mu.$$

The remaining equation on $\mathcal{C}$ is $\Gamma(X^{\rm V}(L)) - X^{\rm C}(L) = 0$. Since $X^{\rm V}(v_z) = 0$ and $X^{\rm C}(v_z) = 0$, both $X^{\rm V}$ and $X^{\rm C}$ are tangent to the constraint, so we can rewrite that equation as $\Gamma(X^{\rm V}(L_c)) - X^{\rm C}(L_c) = 0$, with $L_c = \frac{1}{2}((1+x^2)\dot{x}^2 + \dot{y}^2)$. One easily verifies that this equation is equivalent with

$$(1+x^2)\ddot{x} - x\dot{x}^2 = 0.$$

This equation is evidently $\mathbb{R}^2$-invariant and the reduced vector field is

$$\check{\Gamma} = \dot{x}\frac{\partial}{\partial x} + \frac{x\dot{x}^2}{1+x^2}\frac{\partial}{\partial \dot{x}}.$$

It is instructive to see how one gets the same result when we use Routh reduction. Remark that $H_\mu = \mathbb{R} \times 0$ and $G_\mu = \mathbb{R}^2$. Since also $X^{\rm V}(\dot{y}) = 0$ and $X^{\rm C}(\dot{y}) = 0$ the vector fields $X^{\rm V} = \bar{X}^{\rm V}$ and $X^{\rm C} = \bar{X}^{\rm C}$ are already tangent to the level set $\dot{y} = \mu$. We can therefore simply re-write the remaining equation as

$$\Gamma(X^{\rm V}(\mathcal{R}^c_\mu)) - X^{\rm C}(\mathcal{R}^c_\mu) = 0,$$

where $\mathcal{R}^c_\mu$ is the restriction of the Routhian $\mathcal{R}_\mu$ to the constraints and to the level set. It is given by

$$\mathcal{R}^c_\mu = \tfrac{1}{2}((1+x^2)\dot{x}^2 - \mu^2).$$

The Routh equation above is again $(1+x^2)\ddot{x} - x\dot{x}^2 = 0$. We can now do a direct reduction by means of the largest group G_μ. We see that $N_\mu/G_\mu = T\mathbb{R}$. Due to the absence of gyroscopic-type terms, the G_μ-reduced version of the above equation will be a genuine Euler-Lagrange equation, with the G_μ-reduced Routhian as its Lagrangian. The vector fields $X^{\rm V}$ and $X^{\rm C}$ reduce to the vector fields $\partial/\partial\dot{x}$ and $\partial/\partial x$ on $\mathbb{R}$ and the reduction actually amounts to cancelling the cyclic variables y and z from the above equation. A similar reasoning holds for the reduction in two steps. We have $H_\mu = \mathbb{R}$ and $G_\mu/H_\mu = \mathbb{R}$; the first reduction cancels y and the second cancels z.

Acknowledgements

The first author is a Guest Professor at Ghent University: he is grateful to the Department of Mathematics for its hospitality. The second author is a Postdoctoral Fellow of the Research Foundation – Flanders (FWO).

References

[1] A. M. Bloch with the collaboration of J. Baillieul, P. Crouch and J. E. Marsden, *Nonholonomic Mechanics and Control*, Springer (2003).

[2] A. M. Bloch, P. S. Krishnaprasad, J. E. Marsden and R. M. Murray, Nonholonomic mechanical systems with symmetry, *Arch. Rational Mech. Anal.* **136** (1996) 21–99.

[3] A. M. Bloch, J. E. Marsden and D. V. Zenkov, Quasi-velocities and symmetries in nonholonomic systems, *Dynamical Systems* **24** (2009) 187–222.

[4] F. Cantrijn, J. Cortés, M. de León and D. Martín de Diego, On the geometry of generalized Chaplygin systems, *Math. Proc. Camb. Phil. Soc.* **132** (2002) 323–351.

[5] H. Cendra, J. E. Marsden and T. S. Ratiu, *Lagrangian Reduction by Stages*, Memoirs of the American Mathematical Society 152, AMS (2001).

[6] H. Cendra, J. E. Marsden and T. S. Ratiu, Geometric mechanics, Lagrangian reduction, and nonholonomic systems, In: B. Engquist and W. Schmid (Eds.), *Mathematics Unlimited — 2001 and Beyond*, Springer (2001), 221–273.

[7] J. Cortés Monforte, *Geometric, Control and Numerical Aspects of Nonholonomic Systems*, Lecture Notes in Mathematics 1793, Springer (2002).

[8] M. Crampin and T. Mestdag, Routh's procedure for non-Abelian symmetry groups, *J. Math. Phys.* **49** (2008) 032901.

[9] M. Crampin and T. Mestdag, Reduction and reconstruction aspects of second-order dynamical systems with symmetry, *Acta Appl. Math.* **105** (2009) 241–266.

[10] M. Crampin and T. Mestdag, Anholonomic frames in constrained dynamics, *Dynamical Systems* **25** (2010) 159–187.

[11] M. Crampin and F. A. E. Pirani, *Applicable Differential Geometry*, LMS Lecture Notes 59, Cambridge University Press (1988).

[12] R.H. Cushman, H. Duistermaat and J. Śniatycki, *Geometry of Nonholonomically Constrained Systems*, Advanced Series in Nonlinear Dynamics 26, World Scientific (2010).

[13] M. de León, J. C. Marrero and E. Martínez, Lagrangian submanifolds and dynamics on Lie algebroids, *J. Phys. A: Math. Gen.* **38** (2005) R241–R308.

[14] B. Jovanovic, Geometry and integrability of Euler-Poincaré-Suslov equations, *Nonlinearity* **14** (2001) 1555–1567.

[15] J. Koiller, Reduction of some classical non-holonomic systems with symmetry, *Arch. Rat. Mech. Anal.* **118** (1992) 113–148.

[16] O. Krupková, Mechanical systems with non-holonomic constraints, *J. Math. Phys.* **38** (1997) 5098–5126.

[17] J. E. Marsden and T. S. Ratiu, *Introduction to Mechanics and Symmetry*, Texts in Applied Mathematics 17, Springer (1999).

[18] J.E. Marsden, T.S. Ratiu and J. Scheurle, Reduction theory and the Lagrange-Routh equations, *J. Math. Phys.* **41** (2000) 3379–3429.

[19] T. Mestdag and M. Crampin, Invariant Lagrangians, mechanical connections and the Lagrange-Poincaré equations, *J. Phys. A: Math. Theor.* **41** (2008) 344015.

[20] T. Mestdag and B. Langerock, A Lie algebroid framework for non-holonomic systems, *J. Phys. A: Math. Gen.* **38** (2005) 1097–1111.

[21] J. I. Neĭmark and N. A. Fufaev, *Dynamics of Nonholonomic Systems*, Transl. of Math. Monographs 33, AMS (1972).

[22] W. Sarlet, F. Cantrijn and D. J. Saunders, A geometrical framework for the study of non-holonomic Lagrangian systems, *J. Phys. A: Math. Gen.* **28** (1995) 3253–3268.

[23] J. Vilms, Connections on tangent bundles, *J. Diff. Geom.* **1** (1967) 235–243.